L-022

小型 IoT 機器による適応的ウインドウ法パラメータ推定を用いたネットワークの異常検知
The Methodology of Effective Estimation Parameters Using Adaptive Windowing to Detect Intrusion and Change-Point on Micro Computer System

比賀江 文子† 宮坂 虹規† 嶋 久登†
Fumiko Higae Koki Miyasaka Hisato Shima

1. はじめに

近年、日本国内の製造業は生産性向上を目指し、工場内設備 にセンサを設置しデータの解析結果を用い工場内機器の自動制御や稼働状況の見える化等、様々な用途で産業用 IoT が導入されている。一方、産業用 IoT の導入によって工場内が外部ネットワークに”つながる”ことから、サイバー攻撃等の新たな脅威に対応する必要がある [1]。

情報セキュリティ対策の一環としてネットワーク通信の異常検知が広く用いられているものの不十分といえる [2][3]。工場内ネットワークは特定のセンサや機器との限定的なデータ送受信のみである。そのためデータ構成は極めて単調であるといえる。そこで、この限定的な環境において各種センサやロボットなどから送受信されるパケットを小型 IoT 機器で通信フィルタすることによって、従来手法と比較しより多くの攻撃手法にたいしてリアルタイムに異常検知および変化予兆検知を実現する手法を提案する。本研究におけるシステム概要図を図 1 に示す。

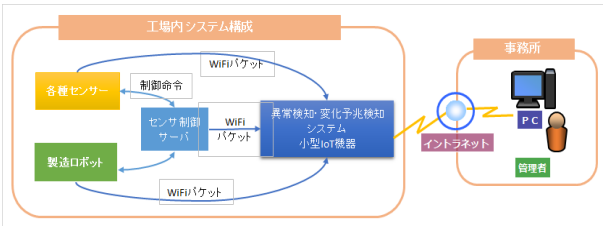

図 1　システム概要図

2. 関連研究

これまでに提案された通信挙動の特徴を用いる手法の多くは、個別の攻撃手法に特化したものであり同時に大規模ネットワーク環境でのパケット単位の情報アクセスが前提となっている。多くの組織ではパケット単位のデータ保有することはコスト的に不可能である。

文献[4]にてリアルタイム MDL 変化統計量に基づいて変化検知を行い、ウインドウの長さをデータに応じて可変とする「適応的ウインドウ法」が提唱されている。非常に有用な手法だが、解析データ量およびモデルとも容量が大きくならざるを得ない。

文献[5]では攻撃者がコマンドサーバおよび制御サーバからネットワーク上の 1 つ以上のホストを感染させる攻撃を検知する手法として、フォレッシング分析における中心性分析を提案している。マルウェアのノードが中心となる性質に着目した特化型分析手法の代表的手法である。

3. システム構成

本研究において提案する本研究では第一次フィルタシステム構成を図 2 に示す。

小型 IoT 機器を用い、工場内各種センサや端末、その制御サーバのパケットを採集し、侵入・ノード追加等パターンマッチングやブラックリストを用いずに不正通信を検知する。

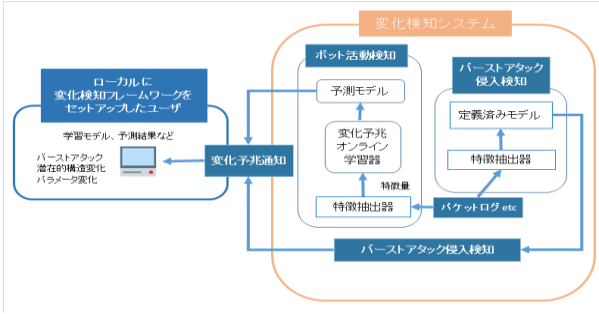

図 2　システム構成図

小型 IoT 機器では容量の大きな学習モデルを搭載させることができない。そこで「適応的ウインドウ法」[4]をさらに簡素化するモデルを考案する。

4. 適応的ウインドウ法のパラメータ推定の手順

「適応的ウインドウ法」を実現させるにはウインドウの長さをデータに応じて可変させる必要がある。推定を始めるにあたってまずデータの可変範囲、平均、中央値、標準偏差、最大値、最小値等、基本統計量を調べる。

本推定の手順を説明する。適切な特徴量を相関分析にて抽出し、その特徴量の残差平均を回帰分析する。検定は残差回帰をカーネル密度推定しウインドウ長(データ送受信間隔)の妥当性を検証する[8]。

4.1 特徴量抽出と回帰分析

相関係数の計算の基準となっている直線が、最小二乗回帰直線であり、2 変量の相関係数、回帰の決定係数、回帰の予測値と被説明変数の相関係数のこれら 3 つの係数が全く一致する[6]。2 変量の場合、説明変数と被説明変数を入れ替えた逆回帰においてもこの 3 つの係数が等しくなるという性質をもつ。そのため、相関係数が強い 2 変量は互いに独立ではないため特徴量として扱えない[6]。

4.2 残差回帰検定

回帰残差正常標本の共役分布を仮定し、異常値を含むデータの標本分布からの比、1 からのずれを測定する[7]。異常値は正常標本分布の比、すなわち密度比が 1 からずれる[8]。この性質を持いて、カーネル密度推定にて残差回帰検定をする[8]。

5. 応的ウインドウ法のパラメータ推定

5.1 特徴量選択の仮説

予備分析の結果、以下のパラメータについて本仮説の検証のため再度相関分析、回帰分析および残差をカーネル密度推定にて検定する。

• パケット送受信時刻
• 送信元 IP アドレス： 送信元ポート番号
• パケットサイズ
• 前回送受信時刻と現在送受信時刻の差分(送信サイクル)

5.2 特徴量選択の仮説検証

1時間単位と1分単位で5.1データを送信元 IP アドレス：送信元ポート番号にて集約し「合計パケットサイズ」「平均パケットサイズ」「パケットサイズの標準偏差」「パケットサイズ最大値」「パケットサイズ最小値」「パケット送信回数」を相関分析した。パケット平均、パケット送信回数、1 時間単位の相関分析のヒートマップと 1 分単位の相関分析のヒートマップに差がないことから、データ毎の「適応的ウインドウ」サイズは 1～60 分間変動するといえる。5.1データについて1時間単位で作成したヒートマップを図 3 に示す。

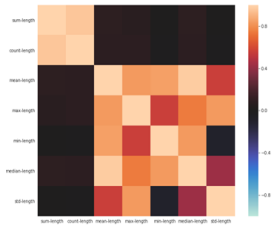

図 3　データ送信間隔とサイズ相関ヒートマップ

5.3 回帰分析と残差回帰検定

5.1 同様データについて回帰分析をする。IP アドレスと Port 番号単位送信サイクルと平均データ量は 1 時間毎および 1 分毎ともに一様分布を得ている。本結果を図 4 にて示す。縦軸に送信サイクル単位での送信サイズの平均、横軸に時間単位内の送信サイクルを示す。以上の結果より送信サイクルと平均データ量は一定サイクルにて一定の量だけ送信させるという規則性があることが分かった。

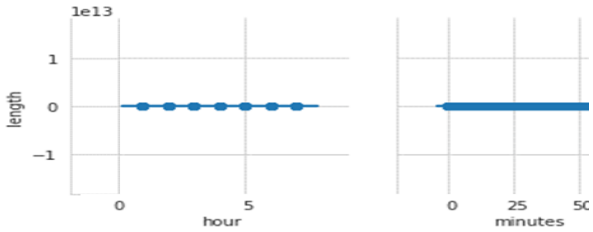

図 4　送信元 IP:Port 毎時および分単位で周期毎とデータ長の単回帰分析

次に残差平均回帰分析の結果をカーネル密度推定にて分布推定による検定をした。その結果を図 5 に示す。

縦軸に送信サイクル単位での送信サイズの平均、横軸に時間単位内の送信サイクルを示す。図 5 より 1 時間単位の送信サイクルにおいては一様分布を得ているものの、1～60 分単位での送信サイクルでは IP アドレス:ポート番号毎に周期およびデータ量は規則性を伴い微妙に変動することが分かった。簡易版「適応的ウインドウ法」は工場内ネットワークという至極限定的な環境の下、用いることを前提にしている。そのため、特定の IP アドレスおよびポートで利用するため、5.1 パラメータはのうちパケット送信間隔を 1 時間単位にてパラメータ推定することで十分事足りるといえる。

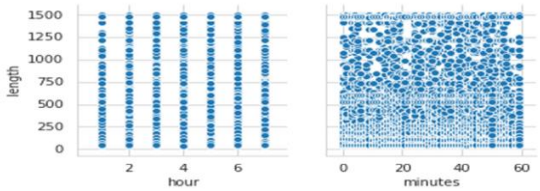

図 5　送信元 IP:Port 毎時間および分単位にて周期毎とデータ長のカーネル密度推定

6. おわりに

極めて限定的環境下における適応性ウインドウ法による異常検知モデルを図 6 に示す。

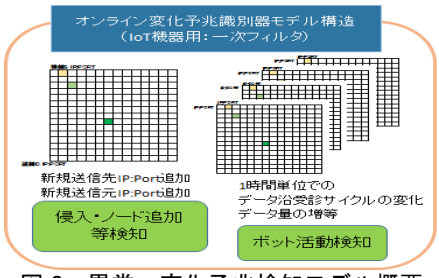

図 6　異常・変化予兆検知モデル概要

時間単位で送信元 IP:Port と送信先 IP:Port 表を保持することでデータ量の平均、サイクルの平均等を表現することができる。例えば IP:PORT が新たに追加された場合、意図的にノードを増やす操作以外はマルウェアのボット活動の可能性があるため、その動作を即時に検知できる。データ送信先の変動および送信サイクルやデータ量が増減時は、攻撃（DDos 攻撃等）によるビジー状態の兆候といえ、開始段階で検知できる。今後、実稼働に向け、監視対象パケットモニタリングおよび変化予兆検知精度を向上させる。

参考文献

[1] “工場における産業用 IoT 導入のための セキュリティ ファーストステップ” 一般社団法人 JPCERT コーディネーションセンター (2018)
[2] 比賀江 文子 “IoT マルウェア Mirai と Hajime の解析”創刊号サイバー犯罪対策総力特集 ZERO DAY 一般財団法人サイバーセキュリティ財団 (2017)
[3] 比賀江 文子 “Shamoon2.0(サウジアラビア空港を襲ったマルウェア) “ 創刊号サイバー犯罪対策総力特集 ZERO DAY 一般財団法人サイバーセキュリティ財団 (2017)
[4] Ryoya Kaneko, Kohei Miyaguchi, Kenji Yamanishi "Detecting changes in streaming data with information-theoretic windowing" IEEE International Conference on Big Data (Big Data) (2017)
[5] Abbas Abou Daya, Mohammad A. Salahuddin, Noura Limam, and Raouf Boutaba "A Graph-Based Machine Learning Approach for Bot Detection" ternational Symposium on Integrated Network Management, Washington DC, USA, April (2019)
[6] 二宮 正司 "相関係数に関する若干の考察" 大阪経済大学論文集 第 59 巻第 4 号 (2008)
[7] 樋田 勉 "カーネル密度推定による適合度検定" 群馬大学社会情報学部研究論集 9, 115-134, (2002)
[8] 谷崎 久志 “密度関数のカーネル推定量におけるバンド幅の選択について： モンテカルロ実験による小標本特性” 国民経済雑誌 第 191 巻, 第 1 号, pp.59-70 (2005)

† 神戸情報大学院大学 Kobe Institute of Computing / Graduate School of Information Technology